

## BLAST Basic Local Alignment Search Tool

Job Title: Nucleotide sequence (15 letters)

---

- Your search parameters were adjusted to search for a short input sequence.
- 
- 

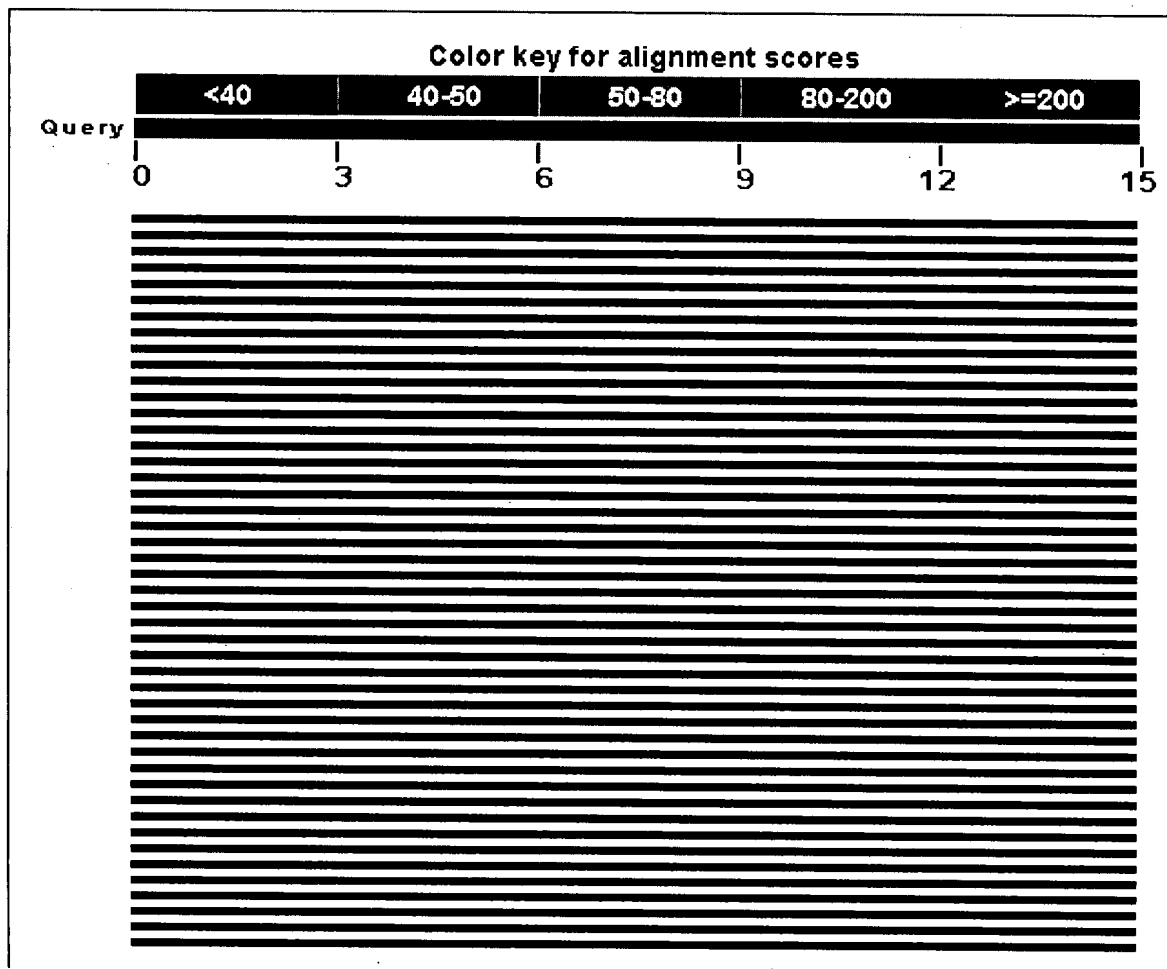
### BLASTN 2.2.17 (Jun-24-2007)

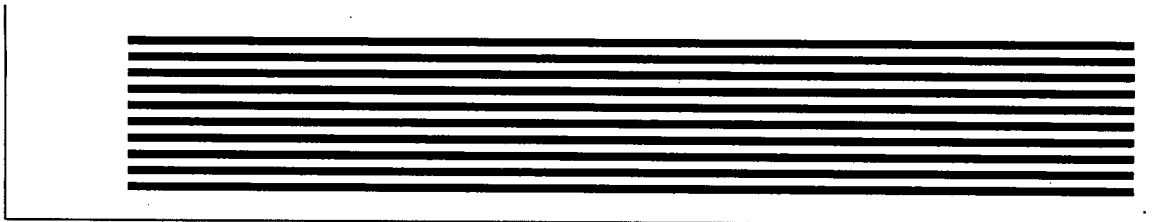
**Reference:** Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402. RID: 8HV76DVM014 **Database:** All GenBank+EMBL+DDBJ+PDB sequences (but no EST, STS, GSS, environmental samples or phase 0, 1 or 2 HTGS sequences) 5,454,516 sequences; 20,739,661,006 total letters

Query= Length=15

### Distribution of 74 Blast Hits on the Query Sequence

---





Distance tree of results **NEW**Legend for links to other resources: **U** UniGene **E** GEO **G** Gene **S** Structure **I****Sequences producing significant alignments:**

(Click headers to sort columns)

<b>AM468046.2</b>	Vitis vinifera contig VV78X022861.5, whole genome shotgun sequence	30.2	30.2	100%	33	100%	
<b>AC202214.3</b>	Pea Aphid NOVECTOR VMRC38-20-A9 () complete sequence	30.2	30.2	100%	33	100%	
<b>NM_018666.2</b>	Homo sapiens sarcoma antigen 1 (SAGE1), mRNA	30.2	30.2	100%	33	100%	<b>UG</b>
<b>CP000660.1</b>	Pyrobaculum arsenaticum DSM 13514, complete genome	30.2	30.2	100%	33	100%	
<b>AC188814.3</b>	Gallus gallus BAC clone CH261-93D11 from chromosome z, complete sequence	30.2	30.2	100%	33	100%	
<b>XM_770878.1</b>	Cryptococcus neoformans var. neoformans B-3501A hypothetical protein (CNBD3780) mRNA, complete cds	30.2	30.2	100%	33	100%	<b>G</b>
<b>AC189953.7</b>	Rhesus Macaque BAC CH250-253F13 () complete sequence	30.2	30.2	100%	33	100%	
<b>CP000569.1</b>	Actinobacillus pleuropneumoniae L20 serotype 5b complete genome	30.2	30.2	100%	33	100%	
<b>XM_001309841.1</b>	Trichomonas vaginalis G3 surface antigen BspA-like (TVAG_272400) mRNA, complete cds	30.2	30.2	100%	33	100%	<b>G</b>
<b>XM_001309840.1</b>	Trichomonas vaginalis G3 surface antigen BspA-like (TVAG_272390) mRNA, complete cds	30.2	30.2	100%	33	100%	<b>G</b>
<b>XM_001304479.1</b>	Trichomonas vaginalis G3 surface antigen BspA-like (TVAG_422450) mRNA, complete cds	30.2	30.2	100%	33	100%	<b>G</b>
<b>AC193609.2</b>	Gallus gallus BAC clone TAM33-18K5 from chromosome z, complete sequence	30.2	30.2	100%	33	100%	
<b>AE016815.3</b>	Ashbya gossypii (= Eremothecium gossypii) ATCC 10895 chromosome II, complete sequence	30.2	30.2	100%	33	100%	
<b>XM_417055.2</b>	PREDICTED: Gallus gallus similar to Chromosome condensation 1-like (LOC418861), mRNA	30.2	30.2	100%	33	100%	<b>UG</b>
<b>CT573001.7</b>	Zebrafish DNA sequence from clone CH211-180A12 in linkage group 12, complete sequence	30.2	30.2	100%	33	100%	
<b>XM_527015.2</b>	PREDICTED: Pan troglodytes cation channel, sperm associated 3 (CATSPER3), mRNA	30.2	30.2	100%	33	100%	<b>G</b>
<b>BC110384.1</b>	Homo sapiens cation channel, sperm associated 3, mRNA (cDNA clone MGC:117279 IMAGE:5166215), complete cds	30.2	30.2	100%	33	100%	<b>UG</b>
<b>BN000272.1</b>	TPA: TPA_exp: Homo sapiens mRNA for CatSper3	30.2	30.2	100%	33	100%	<b>G</b>
<b>CT737238.2</b>	Pan troglodytes chromosome X clone	30.2	30.2	100%	33	100%	

CH251-56J08 map Xq28, complete  
sequence

AC188312.1	Taeniopygia guttata chromosome UNK clone TGMCBa-32H12, complete sequence	30.2	30.2	100%	33	100%	
XR_012260.1	PREDICTED: Macaca mulatta similar to titin isoform N2-A (LOC703527), mRNA	30.2	30.2	100%	33	100%	<b>UG</b>
AC139131.19	Mus musculus chromosome 7, clone RP23-301N20, complete sequence	30.2	30.2	100%	33	100%	
BC101692.1	Homo sapiens cation channel, sperm associated 3, mRNA (cDNA clone MGC:126741 IMAGE:8069198), complete cds	30.2	30.2	100%	33	100%	<b>UG</b>
AC163351.2	Mus musculus BAC clone RP23-150D4 from chromosome 13, complete sequence	30.2	30.2	100%	33	100%	
NM_178019.1	Homo sapiens cation channel, sperm associated 3 (CATSPER3), mRNA	30.2	30.2	100%	33	100%	<b>UE G</b>
AE017344.1	Cryptococcus neoformans var. neoformans JEC21 chromosome 4, complete sequence	30.2	30.2	100%	33	100%	
AC004764.1	Homo sapiens chromosome 5, P1 clone 255g5 (LBNL:H61), complete sequence	30.2	30.2	100%	33	100%	
AC158619.12	Mus musculus 10 BAC RP23-29103 (Roswell Park Cancer Institute (C57BL/6J Female) Mouse BAC Library) complete sequence	30.2	30.2	100%	33	100%	
XM_538633.2	PREDICTED: Canis familiaris similar to cation channel, sperm associated 3 (LOC481512), mRNA	30.2	30.2	100%	33	100%	<b>UG</b>
AC122483.5	Mus musculus BAC clone RP24-366B4 from chromosome 6, complete sequence	30.2	30.2	100%	33	100%	
AC123801.4	Mus musculus BAC clone RP24-374A1 from chromosome 15, complete sequence	30.2	30.2	100%	33	100%	
AC122402.2	Mus musculus BAC clone RP24-108B4 from chromosome 13, complete sequence	30.2	30.2	100%	33	100%	
AC122900.3	Mus musculus BAC clone RP23-39N24 from 15, complete sequence	30.2	30.2	100%	33	100%	
AC164104.5	Mus musculus BAC clone RP23-428G7 from chromosome 6, complete sequence	30.2	30.2	100%	33	100%	
XM_384727.1	Gibberella zeae PH-1 chromosome 2 hypothetical protein (FG04551.1) partial mRNA	30.2	30.2	100%	33	100%	<b>G</b>
AC163759.4	Pan troglodytes BAC clone CH251-258I24 from chromosome unknown, complete sequence	30.2	30.2	100%	33	100%	
AC079851.14	Homo sapiens 12 BAC RP11-349L8 (Roswell Park Cancer Institute Human BAC Library) complete sequence	30.2	30.2	100%	33	100%	
AC011592.5	Homo sapiens chromosome 3, clone RP11-42F12, complete sequence	30.2	30.2	100%	33	100%	<b>E</b>
XM_570454.1	Cryptococcus neoformans var. neoformans JEC21 hypothetical protein (CND02560) partial mRNA	30.2	30.2	100%	33	100%	<b>UG</b>
AC012176.8	Homo sapiens chromosome 16 clone RP11-206F18, complete sequence	30.2	30.2	100%	33	100%	
CR956411.6	Pig DNA sequence from clone CH242-	30.2	30.2	100%	33	100%	

UE  
G

<b>AF274855.3</b>	Homo sapiens chromosome X clone RP11-366F6 map q28, complete sequence	30.2	30.2	100%	33	100%	
<b>U82696.2</b>	Homo sapiens chromosome X clone ICRFXc104-F064, ICRFXc104-G0799, Qc-14E2, Qc-13D8, Qc-7G11, ICRFXc104-E0681, Qc-11C8, Qc-9A9 map q28, complete sequence	30.2	30.2	100%	33	100%	
<b>AC160944.2</b>	Pan troglodytes BAC clone CH251-7D22 from chromosome unknown, complete sequence	30.2	30.2	100%	33	100%	
<b>BX294104.7</b>	Zebrafish DNA sequence from clone CH211-245I22 in linkage group 1, complete sequence	30.2	30.2	100%	33	100%	
<b>NM_208333.1</b>	Ashbya gossypii ATCC 10895 ABR034Wp (AGOS_ABR034W) mRNA, complete cds	30.2	30.2	100%	33	100%	<b>G</b>
<b>AJ278111.1</b>	Homo sapiens mRNA for putative tumor antigen (SAGE gene)	30.2	30.2	100%	33	100%	<b>UE</b> <b>G</b>
<b>AP002799.3</b>	Homo sapiens genomic DNA, chromosome 11q, clone:RP11-806N19, complete sequence	30.2	30.2	100%	33	100%	
<b>BX908786.9</b>	Zebrafish DNA sequence from clone CH211-1308 in linkage group 3, complete sequence	30.2	30.2	100%	33	100%	
<b>AL807770.17</b>	Mouse DNA sequence from clone RP23-206L14 on chromosome 4, complete sequence	30.2	30.2	100%	33	100%	
<b>AL683854.20</b>	Mouse DNA sequence from clone RP23-237K13 on chromosome 4, complete sequence	30.2	30.2	100%	33	100%	
<b>AL591514.7</b>	Mouse DNA sequence from clone RP23-266L14 on chromosome 11, complete sequence	30.2	30.2	100%	33	100%	

## Alignments

---

>emb|AM468046.2| **D** Vitis vinifera contig VV78X022861.5, whole genome shotgun seq  
Length=17913

Score = 30.2 bits (15), Expect = 33  
Identities = 15/15 (100%), Gaps = 0/15 (0%)  
Strand=Plus/Minus

Query 1 GCATTTGTGAAGAGA 15  
|||||  
Sbjct 3572 GCATTTGTGAAGAGA 3558

>gb|AC202214.3| **D** Pea Aphid NOVECTOR VMRC38-20-A9 () complete sequence  
Length=102345

Score = 30.2 bits (15), Expect = 33  
Identities = 15/15 (100%), Gaps = 0/15 (0%)  
Strand=Plus/Minus

Query 1 GCATTTGTGAAGAGA 15  
|||||  
Sbjct 40355 GCATTTGTGAAGAGA 40341

>ref|NM\_018666.2| **UG** Homo sapiens sarcoma antigen 1 (SAGE1), mRNA  
Length=3069

Score = 30.2 bits (15), Expect = 33  
Identities = 15/15 (100%), Gaps = 0/15 (0%)  
Strand=Plus/Plus

Query 1 GCATTTGTGAAGAGA 15  
|||||  
Sbjct 406 GCATTTGTGAAGAGA 420

>gb|CP000660.1| **D** Pyrobaculum arsenaticum DSM 13514, complete genome  
Length=2121076

Features in this part of subject sequence:  
**CRISPR-associated HD domain protein**









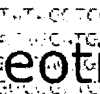
Score = 30.2 bits (15), Expect = 33  
Identities = 15/15 (100%), Gaps = 0/15 (0%)  
Strand=Plus/Plus

Query 1 GCATTTGTGAAGAGA 15  
|||||  
Sbjct 1027001 GCATTTGTGAAGAGA 1027015

>gb|AC188814.3| **D** Gallus gallus BAC clone CH261-93D11 from chromosome z, complet  
sequence  
Length=176534

Score = 30.2 bits (15), Expect = 33  
Identities = 15/15 (100%), Gaps = 0/15 (0%)  
Strand=Plus/Minus

Query 1 GCATTTGTGAAGAGA 15  
|||||  
Sbjct 161540 GCATTTGTGAAGAGA 161526

for

Display  Show   Hide: ☐ sequence ☐ all but gene, CDS and mRNA

Range: from  to  ☐ Reverse complemented strand Features: ☐ SNP

☐ 1: [NM\\_018666](#). Reports Homo sapiens sarc...[gi:145580596]

[Links](#)

[Comment](#) [Features](#) [Sequence](#)

LOCUS NM\_018666 3069 bp mRNA linear PRI 26-JUN-2007  
 DEFINITION Homo sapiens sarcoma antigen 1 (SAGE1), mRNA.  
 ACCESSION NM\_018666  
 VERSION NM\_018666.2 GI:145580596  
 KEYWORDS .  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
 Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3069)  
 AUTHORS Miyahara,Y., Naota,H., Wang,L., Hiasa,A., Goto,M., Watanabe,M.,  
 Kitano,S., Okumura,S., Takemitsu,T., Yuta,A., Majima,Y.,  
 Lemonnier,F.A., Boon,T. and Shiku,H.  
 TITLE Determination of cellularly processed HLA-A2402-restricted novel  
 CTL epitopes derived from two cancer germ line genes, MAGE-A4 and  
 SAGE  
 JOURNAL Clin. Cancer Res. 11 (15), 5581-5589 (2005)  
 PUBMED 16061876  
 REMARK GeneRIF: MAGE-A4(143-151) and SAGE(715-723) are  
 HLA-A2402-restricted CTL epitopes

REFERENCE 2 (bases 1 to 3069)  
 AUTHORS Martelange,V., De Smet,C., De Plaen,E., Lurquin,C. and Boon,T.  
 TITLE Identification on a human sarcoma of two new genes with  
 tumor-specific expression  
 JOURNAL Cancer Res. 60 (14), 3848-3855 (2000)  
 PUBMED 10919659

COMMENT REVIEWED [REFSEQ](#): This record has been curated by NCBI staff. The  
 reference sequence was derived from [AJ278111.1](#), [AL953870.2](#),  
[CD358947.1](#) and [BF056519.1](#).  
 On Apr 20, 2007 this sequence version replaced gi:8924241.

Summary: This gene belongs to a class of genes that are activated  
 in tumors. These genes are expressed in tumors of different  
 histologic types but not in normal tissues, except for  
 spermatogenic cells and, for some, placenta. The proteins encoded  
 by these genes appear to be strictly tumor specific, and hence may  
 be excellent sources of antigens for cancer immunotherapy. This  
 gene is expressed in sarcomas.

COMPLETENESS: complete on the 3' end.

PRIMARY	REFSEQ_SPAN	PRIMARY_IDENTIFIER	PRIMARY_SPAN	COMP
	1-913	AJ278111.1	1-913	
	914-914	AL953870.2	11665-11665	



915-1663	AJ278111.1	915-1663
1664-1951	CD358947.1	502-789
1952-2580	AJ278111.1	1952-2580
2581-3045	BF056519.1	1-465
3046-3069	AJ278111.1	3046-3069

FEATURES

source	Location/Qualifiers
	1..3069
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/chromosome="X"
	/map="Xq26"
gene	1..3069
	/gene="SAGE1"
	/note="sarcoma antigen 1; synonym: SAGE"
	/db_xref="GeneID:55511"
	/db_xref="HGNC:30369"
	/db_xref="HPRD:02290"
	/db_xref="MIM:300359"
CDS	168..2882
	/gene="SAGE1"
	/note="putative tumor antigen"
	/codon_start=1
	/product="sarcoma antigen 1"
	/protein_id="NP_061136.2"
	/db_xref="GI:145580597"
	/db_xref="CCDS:CCDS14652.1"
	/db_xref="GeneID:55511"
	/db_xref="HGNC:30369"
	/db_xref="HPRD:02290"
	/db_xref="MIM:300359"
	/translation="MQASPLQTSQPTPPEELHAAAYVFTNDGQQMRSDEVNLVATGHQ SKKKHSRKSKRHSSSKRRKSMSSWLDKQEDAAVTHSICEERINNGQPVADNVLSTAPP WPDATIAHNIREERMENGQSRTDKVLSTAPPQLVHMAAAGIPSMSTRDLHSTVTHNIR EERMENGQPPQPDNVLSTGPTGLINMAATPIPAMSARDLYATVTHNVCEQKMENVQPAP DNVLLTLRPRRINMTDTGISPMSTRDPYATITYNVPEEKMEKGQQPDNILSTASTGL INVAGAGTPAISTNGLYSTVPHNVCEEKMENDQQPQNNVLSTVQPVIIYLTATGIPGM NTRDQYATITHNVCEERVVNNQPLPSNALSTVLPGLAYLATADMPAMSTRDQHATIIH NLREEKKDNSQPTPDNVLSAVTPELINLAGAGIPPMSTRDQYATVNHVHEARMENGQ RKQDNVLSNVLSGLINMAGASIPAMSSRDLYATITHSVREEKMESGKQPQTDKVISNDA PQLGHMAAGGIPSMSTKDLYATVTQNVHEERMENNQPQPSYDLSTVLPGLTYLTVAGI PAMSTRDQYATVTHNVHEEKIKNGQAASDNVFTVPPAFINMAATGVSSMSTRDQYAA VTHNIREEKINNSQPAPGNILSTAPPWLRHMAAAGISSTITRDLYVTATHSVHEEKMT NGQQAPDNSLSTVPPGCINLSGAGISCRSTRDLYATVIHDIQEEEMENDQTPPDGFLS NSDSPELINMTGHCMPPNALDSFSHDFTSLSKDELLYKPSNEFAVGTKNYSVSAGDP PVTVMISLVETVPNTFQISPAMAKKINDDIKYQLMKEVRRFGQNYERIFILLEEVQGSM KVKRQFVEFTIKEAARFKKVLIQQLKALKEIDSHCHLRKVKHMRKR"
polyA signal	3020..3025
	/gene="SAGE1"
polyA site	3046
	/gene="SAGE1"

ORIGIN

1	caggaccgct	ggcgccagcg	ggagaaccag	gaagcgcagc	tctcccctta	gctgactctg
61	gtgctctcta	caggtggcct	tcactatcaa	ctgcacagcg	ggcagaggca	gaggaacaca
121	gaaccttctc	acacagcggg	ataaaagggt	gaatttggag	ctgacagatg	caggcttctc
181	cacttcaaac	gagtcaacca	actccacctg	aagaacttca	tgctgctgcc	tatgtgttta
241	caaattgatg	gcagcaaatt	aggagtgatg	aagtaaattc	ggttgcaaca	gggcatcaaa
301	gcaaaaagaa	acattccaga	aaatccaaga	gacactcttc	atctaagaga	aggaagagta
361	tgtcctcgtg	gtagacaaa	caggaagatg	ctgcagtcac	tcacagcatt	tgtgaagaga
421	ggataaataa	cggccaacca	gtagctgata	atgtcttgct	aactgctcca	ccatggcctg

481 atgctaccat cgctcacaat atccgtgaag agaggatgga aaatggccaa tctcgaactg  
541 acaaagtctt gtcaactgct ccaccacagc ttgttcatat ggctgcagct ggtattccat  
601 ccatgagtac cagggatctg cattctaccg tcaactcaca tatccgtgaa gagagaatgg  
661 aaaatggcca accccaacct gataacgtct tgtcaactgg tcccacaggg cttattaata  
721 tggcagcaac tcctattcca gccatgagtg ccagagatct ctatgctaca gtcactcaca  
781 atgtctgtga acagaagatg gaaaatgtcc aaccagcacc tgataacgtg ttgttgactc  
841 ttcgaccacg gcgtattaat atgacagaca ctggattttc acccatgagt accagggatc  
901 catatgctac catcacttac aatgtccctg aggagaagat ggaaaagggc caacccaac  
961 ctgataacat cttgtcaact gcttcaacag ggcttattaa tgtggcagga gctggtactc  
1021 cagccatcag caccaatggc ctgtattcca ccgtccctca caatgtctgt gaagagaaga  
1081 tggaaaatga ccaaccgcaa cctaataacg tattgtcaac tgttcaacca gtgattattt  
1141 atttgacagc aactgggtatt ccgggcatga ataaccaggga tcagtatgct accatcactc  
1201 acaatgtctg tgaagagaga gtggtaaata accaaccact acctagtaac gccttgtcaa  
1261 ctgtttctacc agggcttgct tatttggtcaa cagctgatat gccagccatg agtaccaggg  
1321 atcagcatgc taccatcatt cacaatctgc gtgaagagaa gaaagataac agccaacca  
1381 cccctgataa cgtcttgtca gctgttacac cagagcttat taacttggca ggagctggtg  
1441 ttccacccat gagtaccagg gatcagtatg ctaccgtcaa tcaccatgtc catgaagcaa  
1501 ggatggaaaa tggccaacga aaacaggata acgtcttgtc aaatgttcta tccgggctta  
1561 ttaatatggc aggagctagt attccagcaa tgagttccag ggatctgtat gctaccatta  
1621 ctacagtggt tcgtgaagag aagatggaaa gtggcaaacc ccaaactgat aaggtcatat  
1681 caaatgatgc accacagctt ggtcatatgg ctgcaggtgg tattccatcc atgagtacca  
1741 aggatctgta tgctaccgtc actcaaatg tccatgaaga gaggatggaa aataaccaac  
1801 cacaacctag ttatgacttg tcaactgttc taccaggact tacttatttg acagtagctg  
1861 gtattccggc catgagtacc agggatcagt atgctaccgt cactcacaat gtccatgaag  
1921 agaagattaa aaatggccaa gcagcatccg ataatgtctt ctcgactgtt ccaccagcat  
1981 ttattaatat ggcagcaact ggtgtttcat ccatgagtac cagggatcag tatgctgcag  
2041 tcaactcaca catccgtgaa gagaagataa ataacagcca accagcacct ggtaacatct  
2101 tgtcaactgc tcctccatgg cttcgtcata tggcagcagc tgggaatttca tccacgatta  
2161 ccagggatct gtatgtcacc gccactcaca gtgtccatga ggagaagatg acaaatggcc  
2221 aacaggcacc tgataactcc ttgtcaacgg ttccacctgg ttgtattaat ctgtcaggag  
2281 ctgggtatttc atgcagaagt accagggatc tgtatgctac tgtcattcac gatatccagg  
2341 aggaggagat ggaaaatgat caaacccctc ctgatggctt cctgtcaaat tctgattcac  
2401 cagagctgat aaatatgaca ggacattgta tgccacccaa tgcattggat tctttctctc  
2461 acgacttcac aagtctcagc aaagatgagc tgctttacaa acctgatagt aatgaatttg  
2521 cggtaggcac caaaaactac agtgtctctg caggtgaccc accagttaca gtaatgtctt  
2581 tgggtggaaac tgtgccaaat acaccacaaa tatctcctgc catggcaaag aaaattaatg  
2641 atgatataaa atatcaatta atgaaagaag ttcgaaagggt tgggcaaaat tatgaaagaa  
2701 ttttcatatt gcttgaagag gtacaaggat ctatgaaagt caagagacaa tttgttgaat  
2761 ttaccatcaa ggaagcagca aggtttaaaa aagttgtctt aattcagcaa ctcgagaagg  
2821 cgcttaaaaga aatagattcc cactgccatc tcagaaaagt taagcacatg agaaaaagat  
2881 aattgtgtta gtgcaaagac caaggagaaa caaggacata tgctgtagga tggacagggt  
2941 tattgctgaa gctccctata atcctgaaat gaagagaatt cccttcaga agctacgaaa  
3001 aaggagctg tttaaattta ataaatctct gttagtaaaa gctgcaaaaa aaaaaaaaaa  
3061 aaaaaaaaaa

//

[Disclaimer](#) | [Write to the Help Desk](#)  
[NCBI](#) | [NLM](#) | [NIH](#)

Jun 19 2007 13:56:00